Nucleic Acids Research, Vol. 19, No. 19 5233–5236

Orientation of the Lac repressor DNA binding domain in complex with the left *lac* operator half site characterized by affinity cleaving

Jumi A.Shin, Richard H.Ebright[1] and Peter B.Dervan*
Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125 and [1]Department of Chemistry and Waksman Institute, Rutgers University, New Brunswick, NJ 08855, USA

Received July 19, 1991; Revised and Accepted August 22, 1991

ABSTRACT

Lac repressor (LacR) is a helix-turn-helix motif sequence-specific DNA binding protein. Based on proton NMR spectroscopic investigations, Kaptein and co-workers have proposed that the helix-turn-helix motif of LacR binds to DNA in an orientation opposite to that of the helix-turn-helix motifs of λ repressor, λ cro, 434 repressor, 434 cro, and CAP [Boelens, R., Scheek, R., van Boom, J. and Kaptein, R., *J. Mol. Biol.* 193, *1987*, 213–216]. In the present work, we have determined the orientation of the helix-turn-helix motif of LacR in the LacR-DNA complex by the affinity cleaving method. The DNA cleaving moiety EDTA·Fe was attached to the N-terminus of a 56-residue synthetic protein corresponding to the DNA binding domain of LacR. We have formed the complex between the modified protein and the left DNA half site for LacR. The locations of the resulting DNA cleavage positions relative to the left DNA half site provide strong support for the proposal of Kaptein and co-workers.

INTRODUCTION

Escherichia coli Lac repressor (LacR) is a transcriptional regulatory protein; LacR functions by binding, in the absence of the inducer molecules allolactose or isopropyl-β-*D*-thiogalactoside, to a specific DNA site, *lac* operator, located adjacent to the *lac* promoter (reviewed in ref. 1). LacR and its specific DNA site have been subjected to extensive structural, biophysical, biochemical, and genetic analyses (2–5). LacR is a tetramer of four chemically identical subunits, each consisting of 360 amino acids. The 56 N-terminal amino acids constitute the DNA binding domain of LacR. The DNA binding domain of LacR can be cleaved proteolytically from the remainder of the molecule to yield a 56-residue fragment that folds independently and that binds to the specific DNA site (6–8). The three-dimensional structure of the DNA binding domain of LacR has been determined by proton NMR spectroscopic investigations (9). The consensus DNA site for LacR is 5′-AATTGTGAGCGCTCACAATT-3′ (10,11); the site is 20 base pairs in length and exhibits perfect two-fold sequence symmetry. The left DNA half site for LacR is 5′-AATTGTGAGC-3′.

LacR is a member of a class of at least 150 sequence-specific DNA binding proteins—prokaryotic and eukaryotic—that utilize a highly conserved α-helix-turn-α-helix motif to mediate interaction with DNA (12–14; reviewed in refs. 15,16). Five other members of this class (i.e., λ repressor [17], λ cro [18], 434 repressor [19], 434 cro [20], and CAP [21]) have been documented to interact with DNA as illustrated schematically in Figure 1A. In the case of LacR, amino acids 5 through 25 constitute the helix-turn-helix motif.

Two models have been proposed for the structure of the LacR-DNA complex. The models differ in the orientation of the helix-turn-helix motif of each subunit with respect to the DNA site. Based on amino acid sequence and three-dimensional structural homology considerations, Matthews and co-workers (12; see also ref. 22) proposed that the helix-turn-helix motif-DNA orientation in the case of LacR is similar or identical to the orientation documented in the cases of λ repressor (17), λ cro (18), 434 repressor (19), 434 cro (20), and CAP (21) (Figure 1A). In this model, the N-terminus of the helix-turn-helix motif of LacR is near to base pair 4 of the DNA half site, and the C-terminus of the helix-turn-helix motif of LacR is near to base pair 10 of the DNA half site. In contrast, based on proton NMR spectroscopic investigations, Kaptein and co-workers (23–26) have proposed that the helix-turn-helix motif-DNA orientation in the case of LacR is opposite—180° different—from the orientation in the cases of the other proteins (Figure 1B). In this model, the N-terminus of the helix-turn-helix motif of LacR is near to base pair 10 of the DNA half site, and the C-terminus of the helix-turn-helix motif of LacR is near to base pair 4 of the DNA half site.

Recent genetic (27) and photocrosslinking (28) results provide strong support for the model of Kaptein and co-workers (23–26; Figure 1B). Genetic results suggesting that amino acid 18 of LacR

* To whom correspondence should be addressed

contacts base pair 7 of the DNA half site (22,29,30) are consistent both with the model of Matthews and co-workers (12) and with the model of Kaptein and co-workers (23–26). Genetic results suggesting that amino acid 22 of LacR contacts base pair 5 of the DNA half site (27) are consistent only with the model of Kaptein and co-workers (23–26). Photocrosslinking results suggest that amino acid 17 of LacR is close to base pair 8 of the DNA half site, and that amino acid 29 of LacR is close to base pairs 3 and 4 of the DNA half site (28). In the present work, we have obtained additional strong support for the model of Kaptein and co-workers (23–26).

The DNA cleaving moiety, EDTA·Fe, can be incorporated at discrete amino acid residues in a protein, thus allowing the positions of these residues in a DNA complex to be mapped to nucleotide resolution by high resolution gel electrophoresis (31–37). We have used the affinity cleaving method to determine the location of the N-terminus of LacR within the LacR-DNA complex. We have attached EDTA·Fe to the N-terminus of the 56-residue LacR DNA binding domain peptide LacR(1–56), to construct [(EDTA·Fe)-γ-aminobutyryl]LacR(1–56) (Figure 2). After formation of the [(EDTA·Fe)-γ-aminobutyryl]LacR (1–56)-DNA complex, DNA cleavage is initiated by addition of the reducing agent sodium ascorbate. We have found that DNA cleavage occurs just beyond base pair 10 of the DNA half site.

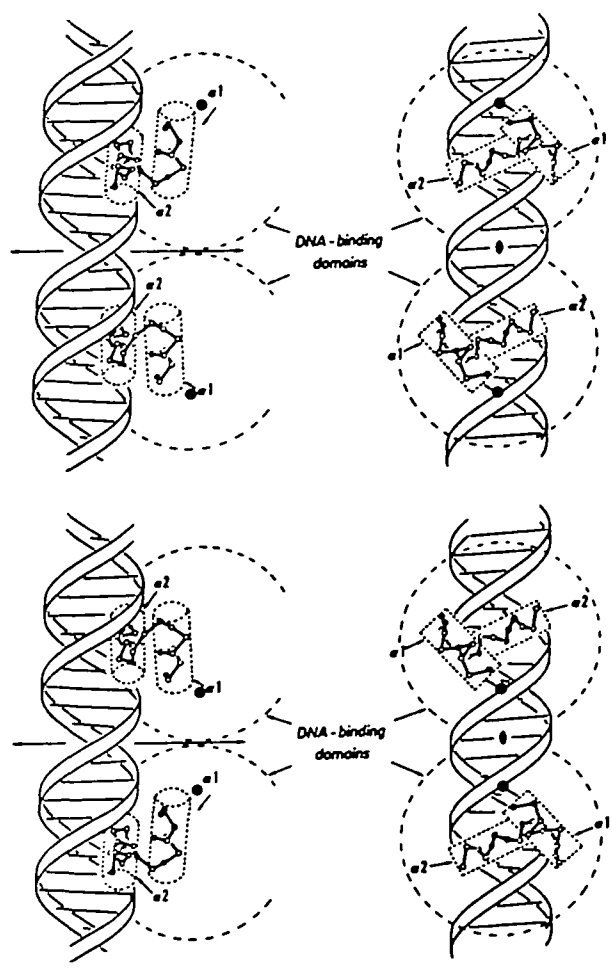

Figure 1. The two models for the structure of the LacR-DNA complex. Heavy filled circles indicate the approximate position of the N-terminus of each subunit. **A.** Model of Matthews and co-workers (12; see also ref. 22); this model is based on the experimentally documented models for the structures of the protein-DNA complexes of λ repressor (17), λ cro (18), 434 repressor (19), 434 cro (20), and CAP (21). The helix-turn-helix motifs of two subunits of the LacR tetramer are proposed to make equivalent, two-fold related contacts with adjacent DNA major grooves. The second α-helix of each helix-turn-helix motif (the 'recognition α-helix') is proposed to interact within the major groove. **B.** Model of Kaptein and co-workers (23–26); the orientation of the helix-turn-helix motif of each subunit is inverted relative to the orientation in A.

MATERIALS AND METHODS

[(EDTA·Fe)-γ-aminobutyryl]LacR(1–56)

Manual peptide synthesis was carried out by solid phase techniques in 20 ml vessels fitted with coarse glass frits, using synthetic protocols developed at the California Institute of Technology (38–40). Fully-protected, resin-bound LacR(1–56) was synthesized on benzhydrylamine resin using t-Boc protected amino acids. γ-aminobutyric acid was incorporated at the deprotected N-terminus of otherwise-protected resin-bound LacR(1–56) using t-butoxycarbonyl-γ-aminobutyric acid. EDTA was incorporated at the deprotected N-terminus of otherwise protected resin-bound [γ-aminobutyric acid]LacR(1–56) using tricyclohexyl-EDTA (41) and standard coupling chemistry. [EDTA-γ-aminobutyryl]LacR(1–56) was cleaved from the resin and deprotected using anhydrous hydrofluoric acid in the presence of p-cresol and p-thiocresol radical scavengers.

[EDTA-γ-aminobutyryl]LacR(1–56) was purified by reverse-phase HPLC on a semipreparative C_8 column (Vydac) with a linear gradient of acetonitrile-water with 0.1% trifluoroacetic acid (flow rate, 3ml/min, 0 to 60% acetonitrile over 240 min). The peptide was homogeneous by the criteria of HPLC, amino acid analysis, and amino acid sequencing. Mass spectrometric analysis: calculated, 6584.5; found 6587±13. [EDTA-γ-aminobutyryl]LacR(1–56) was stored dry at −70°C in 5 nmol aliquots (ϵ_{275}=5620 for four Tyr residues). [(EDTA·Fe)-γ-aminobutyryl]LacR(1–56) was prepared by incubation of [EDTA-γ-aminobutyryl]LacR(1–56) with aqueous ferrous ammonium sulfate in a 1:1 molar mixture (30 min, 25°C).

DNA substrate

Plasmid pJS18H is a derivative of pUC18 (41) having a 66 base pair *Xma*I-*Pst*I insert containing the left DNA half site for LacR, 5′-AATTGTGAGC-3′. Plasmid pJS18H was linearized by digestion with *Eco*RI. Linearized plasmid pJS18H was 3′-end-labeled with [α^{32}P]-dATP and DNA polymerase I Klenow fragment (42). Linearized plasmid pJS18H was 5′-end-labeled by dephosphorylation using calf intestinal alkaline phosphatase, followed by phosphorylation using [γ^{32}P]-ATP and T4 polynucleotide kinase (42). Labeled linearized plasmid pJS18H was digested with *Nde*I, and the resulting labeled 317 base pair DNA fragment was isolated using non-denaturing polyacrylamide gel electrophoresis.

DNA cleavage experiments

Reaction mixtures (10μl) contained ^{32}P end-labeled DNA fragment (20,000 cpm), 50 μM [(EDTA·Fe)-γ-aminobutyryl]LacR (1–56), 1 mM sodium ascorbate, 50 mM tris-acetate, pH 7.0, 20 mM NaCl, 0.1 mg/ml tRNA (Sigma Chemical, Type XX), and 0.5 mg/ml bovine serum albumin. All components except sodium ascorbate were incubated for 30 min at 37°C. Reactions were initiated by the addition of sodium ascorbate (1 mM), and

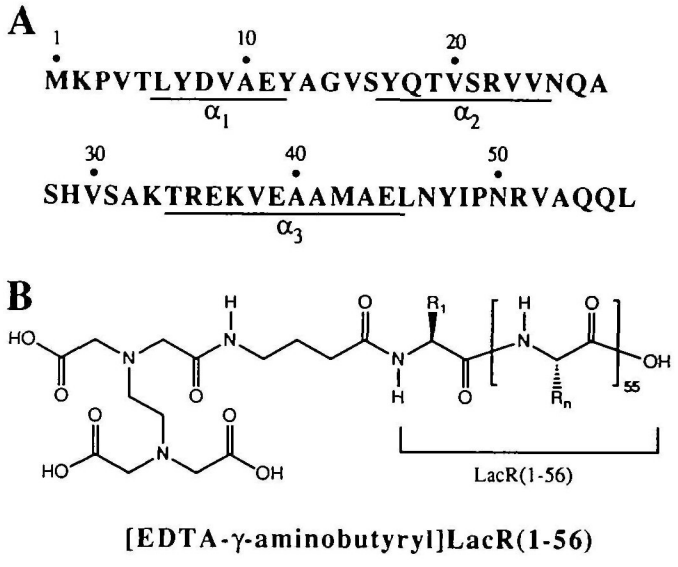

Figure 2. A. The sequence of amino acids 1 to 56 of LacR. **B.** Synthetic protein [EDTA-γ-aminobutyryl]LacR(1−56).

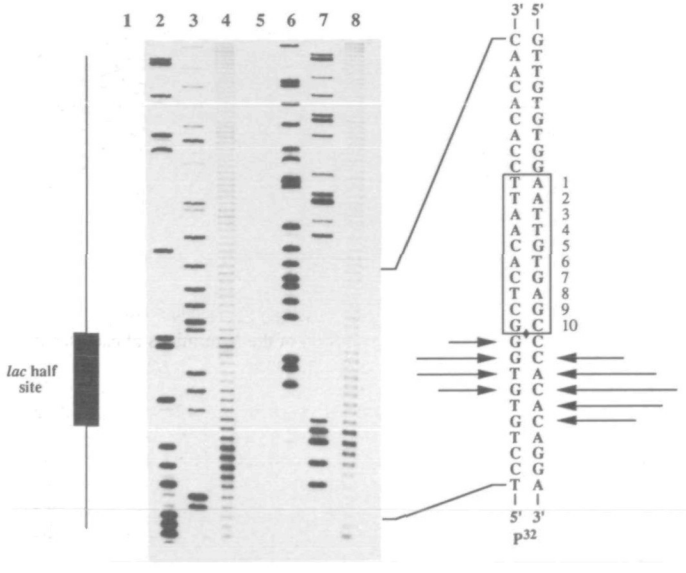

Figure 3. *left panel.* Autoradiogram of a high-resolution denaturing polyacrylamide gel. Lanes 1−4 present data for DNA 3′-end-labeled on the right DNA strand. Lanes 5−8 present data for DNA 5′-end-labeled on the left DNA strand. Lanes 1 and 5, intact DNA; lanes 2 and 6, chemical-sequencing A reaction (46); lanes 3 and 7, chemical sequencing G reaction (42); lanes 4 and 8, [(EDTA·Fe)-γ-aminobutyryl]LacR(1−56) affinity cleavage. *right panel.* [(EDTA·Fe)-γ-aminobutyryl]LacR(1−56) affinity cleavage. The figure illustrates the 10 base pair left DNA half site for LacR. Arrows indicate nucleotides at which DNA cleavage above-background is observed; length of the arrows indicates the relative extent of cleavage.

were allowed to proceed for 30 min at 22°C. Reactions were terminated by desalting using Sephadex G-50 spin columns. Reaction products were analyzed by denaturing electrophoresis on 8% polyacrylamide gels (1:20 crosslink, 7M urea) (43). After electrophoresis, gels were dried and autoradiographed. Autoradiograms were analyzed by laser densitometry.

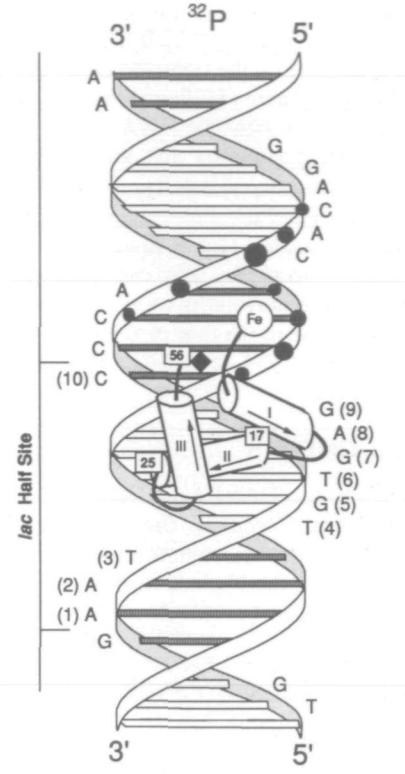

Figure 4. Model for the structure of the complex between [(EDTA·Fe)-γ-aminobutyryl]LacR(1−56) and the left DNA half site. α-Helices are illustrated as cylinders with arrows pointing from the N-terminus to the C-terminus. Filled circles indicate nucleotides at which DNA cleavage is observed; the diameter of the circles indicates the relative extent of cleavage. The model is in agreement with the model of Kaptein and co-workers (23−26; Figure 1B).

RESULTS AND DISCUSSION

In the DNA cleavage experiments reported here, we have investigated the left DNA half site for LacR, rather than the full, two-fold symmetric consensus DNA site for LacR. A single cleavage pattern resulting from one bound EDTA·Fe moiety would result in the least ambiguous data with regard to orientation assignment. Parallel DNA cleavage experiments with the full two-fold symmetric consensus DNA site for LacR yield consistent results.

DNA cleavage experiments were performed with [(EDTA·Fe)-γ-aminobutyryl]LacR(1−56) and a 317 base pair DNA fragment containing the left DNA half site in the presence of the reducing agent sodium ascorbate. DNA cleavage significantly above-background level was observed at five adjacent nucleotides on one DNA strand and at four adjacent nucleotides on the opposite DNA strand (Figure 3). These nucleotide positions are located immediately beyond base pair 10 of the left DNA half site: i.e., at base pairs 11, 12, 13, 14, 15 and 16. These results establish that the EDTA·Fe moiety of [(EDTA·Fe)-γ-aminobutyryl]LacR (1−56) is near to base pairs 11 to 16 of the left DNA half site in the [(EDTA·Fe)-γ-aminobutyryl]LacR(1−56)-DNA complex. These results cannot easily be reconciled with the model of Matthews and co-workers (12; see also ref. 22, Figure 1A). In contrast, these results are in good agreement with the model of Kaptein and co-workers (23−26; Figures 1B,4).

The five adjacent nucleotides at which above-background DNA cleavage was observed on one DNA strand are shifted

approximately one nucleotide in the 3′ direction relative to the four adjacent nucleotides at which above-background DNA cleavage was observed on the opposite DNA strand (Figure 3). This DNA cleavage pattern is diagnostic of DNA cleavage by EDTA·Fe located in or near the DNA minor groove (31–37). Therefore, this result suggests that the EDTA·Fe moiety of [(EDTA·Fe)-γ-aminobutyryl]LacR(1–56) is near the DNA minor groove in the [(EDTA·Fe)-γ-aminobutyryl]LacR(1–56)-DNA complex (Figure 4). This is in agreement with the model of Kaptein and co-workers. We conclude, as proposed by Kaptein, that the helix-turn-helix motif of LacR binds to DNA in an orientation opposite to that of the helix-turn-helix motifs of λ repressor, λ cro, 434 repressor, 434 cro, and CAP.

ACKNOWLEDGMENTS

We are grateful to the DARPA University Research Initiative Program, Merck and Company, the Searle Scholars Program, and the National Science Foundation for support of this research. We are also grateful to the National Institute of General Medical Sciences for a Research Service Award to J.A.S.

REFERENCES

1. Jacob,F. and Monod,J. (1961) *J. Mol. Biol.*, **3**, 318–356.
2. Barkley,M. and Bourgeois,S. (1978) In Miller,J. and Reznikoff,W. (eds.), The Operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, pp. 177–220.
3. Beyreuther,K. (1978) In Miller,J. and Reznikoff,W. (eds.), The Operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, pp. 123–154.
4. Miller,J. (1978) In Miller,J. and Reznikoff,W. (eds.), The Operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, pp. 31–88.
5. Weber,K. and Geisler,N. (1978) In Miller,J. and Reznikoff,W. (eds.), The Operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, pp. 155–176.
6. Geisler,N. and Weber,K. (1978) *FEBS Lett.* **87**, 215–218.
7. Arndt,K., Boschelli,F., Lu,P. and Miller,J. (1981) *Biochem.* **20**, 6109–6118.
8. Nick,H., Arndt,K., Boschelli,F., Jarema,M.A., Lillis,M., Sadler,J., Caruthers,M. and Lu,P. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 218–222.
9. Kaptein,R., Zuiderweg,E., Scheek,R., Boelens,R. and von Gunsteren,W. (1985) *J. Mol. Biol.* **182**, 179–182.
10. Sadler,J., Sasmor,H. and Betz,J. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 6785–6789.
11. Simons,A., Tils,D., von Wilcken-Bergmann,B. and Müeller-Hill,B. (1984) *Proc. Natl. Acad. Sci. USA* 81, 1624–1628.
12. Matthews,B., Ohlendorf,D., Anderson,W. and Takeda,Y. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 1428–1432.
13. Zuiderweg,E., Kaptein,R. and Wüthrich,K. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 5837–5841.
14. Zuiderweg,E., Billeter,M., Boelens,R., Scheek,R., Wüthrich,K. and Kaptein,R. (1984) *FEBS Lett.* **174**, 243–247.
15. Pabo,C. and Sauer,R. (1984) *Ann. Rev. Biochem.* **53**, 293–321.
16. Takeda,Y., Ohlendorf,D., Anderson,W. and Matthews,B. (1983) *Science* **221**, 1020–1026.
17. Jordan,S. and Pabo,C. (1988)*Science* **242**, 893–899.
18. Brennan,R., Roderick,S., Takeda,Y. and Matthews,B. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 8165–8169.
19. Aggarwal,A., Rodgers,D., Drottar,M., Ptashne,M. and Harrison,S. (1988) *Science* **242**, 899–907.
20. Mondragon,A. and Harrison,S. (1990) *J. Mol. Biol.* **219**, 321–334.
21. Steitz,T. (1990) *Quart. Rev. Biophys.* **23**, 205–280.
22. Ebright,R. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 303–307.
23. Boelens,R., Scheek,R., van Boom,J. and Kaptein,R. (1987) *J. Mol. Biol.* **193**, 213–216.
24. Lamerichs,R., Boelens,R., van der Marel,G., van Boom,J., Kaptein,R., Buck,F., Fera,B. and Rüterjans,H. (1989) *Biochem.* **28**, 2985–2991.
25. Kaptein,R., Lamerichs,R., Boelens,R. and Rullmann,J.A. (1990) *Biochem. Pharm.* **40**, 89–96.
26. Lamerichs,R., Boelens,R., van der Marel,G., van Boom,J. and Kaptein,R. (1990) *Eur. J. Biochem.* **194**, 629–637.
27. Lehming,N., Sartorius,J., Oehler,S., von Wilcken-Bergmann,B. and Müller-Hill,B. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7947–7951.
28. Allen,T., Wick,K. and Matthews,K. (1991) *J. Biol. Chem.* **266**, 6113–6119.
29. Ebright,R. (1985) *J. Biomol. Struct. Dyn.* **3**, 281–297.
30. Ebright,R. (1991) *Methods in Enzymol.*, in press.
31. Taylor,J., Schultz,P. and Dervan,P. (1984) *Tetrahedron* **40**, 457–465.
32. Dervan,P. (1986) *Science* **232**, 464–471.
33. Sluka,J., Horvath,S., Bruist,M., Simon,M. and Dervan,P. (1987) *Science* **238**, 1129–1132.
34. Mack,D., Sluka,J., Shin,J., Griffin,J., Simon,M. and Dervan,P. (1990) *Biochem.* **29**, 6561–6567.
35. Graham,K. and Dervan,P. (1990) *J. Biol. Chem.* **27**, 16534–16540.
36. Oakley,M. and Dervan,P. (1990) *Science* **248**, 847–850.
37. Dervan,P. (1991) *Methods in Enzymol..*, in press.
38. Kent,S. (1988) *Ann. Rev. Biochem.* **57**, 957.
39. Clark-Lewis,O., Aebersold,R., Ziltener,H., Schrader,J., Hood,L. and Kent,S. (1986) *Science* **231**, 134–139.
40. Sluka,J., Griffin,J., Mack,D. and Dervan,P. (1990) *J. Am. Chem. Soc.* **112**, 6369–6374.
41. Yanisch-Perron,C., Viera,J. and Messing,J. (1985) *Gene* **33**, 103–119.
42. Maxam,A. and Gilbert,W. (1980) *Methods in Enzymol.* **65**, 499–560.
43. Ogata,R. and Gilbert,W. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5851–5854.
44. Ogata,R. and Gilbert,W. (1979) *J. Mol. Biol.* **132**, 709–728.
45. Sasmor,H. and Betz,J. (1990) *Gene* **89**, 1–6.
46. Iverson,B. and Dervan,P. (1987) *Nucleic Acids Res.* **15**, 7823.